ABSTRACT

A spell checker based on the noisy channel model has a source model and an error model. The source model determines how likely a word w in a dictionary is to have been generated. The error model determines how likely the word w was to have been incorrectly entered as the string s (e.g., mistyped or incorrectly interpreted by a speech recognition system) according to the probabilities of string-to-string edits. The string-to-string edits allow conversion of one arbitrary length character sequence to another arbitrary length character sequence.